Title: Natural diversity of telomere length distributions across 100 yeast strains

Authors: Clotilde Garrido[1,2], Etienne Kornobis[3,4], Nicolas Agier[1,2], Oana Ilioaia[1,2], Gilles Fischer[1,2,*], Zhou Xu[1,2,*]

Affiliations:

[1] Sorbonne Université, CNRS, Laboratory of Computational, Quantitative and Synthetic Biology, CQSB, F-75005 Paris, France.

[2] Sorbonne Université, CNRS, Inserm, Institut de Biologie Paris-Seine, IBPS, F-75005 Paris, France.

[3] Institut Pasteur, Université Paris Cité, Plate-forme Technologique Biomics, F-75015 Paris, France.

[4] Institut Pasteur, Université Paris Cité, Bioinformatics and Biostatistics Hub, F-75015 Paris, France.

Correspondence:

gilles.fischer@sorbonne-universite.fr

zhou.xu@sorbonne-universite.fr

Keywords: telomere, yeast, natural isolates, long-read sequencing

Abstract

Telomerase counteracts the gradual shortening of telomeres that is occurring at each cell division by elongating the telomeric sequences. This process generates a dynamic balance between elongation and shortening, best represented by a steady-state telomere length (TL) distribution. We developed a computational pipeline for detecting telomeric sequences in *Saccharomyces cerevisiae* genome directly from raw Oxford Nanopore reads. This method provides a complete view of TL distributions and enables broad exploration of TL diversity, as it is directly applicable to any high-throughput single-molecule sequencing dataset without prior DNA preparation. We analyzed the TL distribution in 100 *S. cerevisiae* strains representing the genetic and ecological diversity of the species. We obtained the first species-level view of TL distributions at individual chromosome extremities and measured their mean value across all strains at 359 ± 129 bp. Our analysis revealed a large diversity in TL distributions across strains, polyploids exhibiting longer telomeres than diploids and haploids. We confirmed that TL distributions are specific to each chromosome extremity and found that the extent of the inter-extremity differences within a strain varied widely across strains. Each strain displays a distinct TL distribution pattern defined by two components: inter-extremity differences in mean TL and within-extremity TL heterogeneity. Differences in inter-extremity mean TL accounted for most of the observed TL variations between strains. Furthermore, we found that inter-extremity differences partly result from the presence of subtelomeric Y' elements, which are associated with longer telomeres in cis. Interestingly, despite large TL variations, the length of the shortest telomeres remained relatively constant across strains and fluctuated around 138 ± 45 bp, suggesting a selective constraint at the species level. This study reveals the species-wide TL diversity in *S. cerevisiae* and highlights key factors shaping TL distributions at individual chromosome extremities.

Introduction

Telomeres are repeated sequences found at chromosome extremities, which are essential for genome integrity and control of cell proliferation (de Lange, 2018; Jain and Cooper, 2010; Wellinger and Zakian, 2012). These functions are tightly associated with the regulation of telomere length (TL). Indeed, telomeres are protected by proteins that prevent natural chromosome ends from being recognized as double-strand breaks, which would elicit a DNA damage response and inappropriate repair. In the budding yeast *Saccharomyces cerevisiae*, these include Rap1 and its cofactors Rif1/Rif2 bound to double-stranded DNA (dsDNA) in numbers roughly proportional to TL (Levy and Blackburn, 2004; Marcand et al., 1997). Besides, due to the incomplete replication of chromosome extremities, telomeres shorten at each cell division until they reach a critical length that triggers the DNA damage checkpoint and replicative senescence, a programmed cell cycle arrest that prevents further proliferation (Enomoto et al., 2002; Ijpma and Greider, 2003; Lundblad and Szostak, 1989). TL must therefore maintain a minimal length to allow cell division. In most species, including budding yeast, this is done by expressing the telomerase enzyme, which adds telomeric repeats preferentially on shorter telomeres (Greider and Blackburn, 1985; Lundblad and Szostak, 1989; Teixeira et al., 2004). In *S. cerevisiae*, telomerase is constitutively expressed, which leads to a complex dynamics of progressive telomere shortening interrupted by stochastic elongation by limiting amounts of telomerase acting a few telomeres per cell cycle (Gallardo et al., 2011; Mozdy and Cech, 2006; Xu et al., 2013). This length regulation ensures an overall equilibrium, but individual TLs are very variable across different extremities in a given cell and also across different cells for a given extremity (Shampay and Blackburn, 1988). TL is thus best described by a steady-state TL distribution and each species appears to have it set at a characteristic length range (Blackburn and Chiou, 1981; Cristofari and Lingner, 2006; Eberhard et al., 2019; Fajkus et al., 1995; Gomes et al., 2011; Monaghan et al., 2018; Moyzis et al., 1988; Shakirov and Shippen, 2004; Walmsley and Petes, 1985; Wicky et al., 1996).

Despite this steady-state TL distribution, the average TL varies broadly across individuals and subpopulations within a species (Allsopp et al., 1992; D'Angiolo et al., 2023; Eberhard et al., 2019; Fulcher et al., 2015; Gatbonton et al., 2006; Karimian et al., 2024; Liti et al., 2009; O'Donnell et al., 2023; Shakirov and Shippen, 2004; Walmsley and Petes, 1985). In *S. cerevisiae*, early studies based on terminal restriction fragment Southern blots revealed moderate variations in the average TL in a limited number of laboratory and wild strains (Gatbonton et al., 2006; Liti et al., 2009; Walmsley and Petes, 1985). More recently, larger variations in TL across natural isolates were uncovered (D'Angiolo et al., 2023; O'Donnell et al., 2023). However, these variations remain poorly characterized since until recently, it was difficult to experimentally access the full TL distribution. For instance, Southern blotting of the telomere-containing restriction fragments have been the standard method for assessing TL in yeast but does not allow the accurate measurement of TL distribution globally and at each chromosome extremity. Recently, Sholes and colleagues used Oxford Nanopore sequencing to measure the full TL distribution and reported that in the W303 laboratory strain, each chromosome end had a specific and stable TL distribution (Sholes et al., 2022). We previously generated the telomere-to-telomere genome assemblies for 142 natural yeast isolates, which allowed us to also detect extremity-specific TLs (O'Donnell et al., 2023). These observations, since then extended to human telomeres (Karimian et al., 2024; Sanchez et al., 2024; Schmidt et al., 2024), suggested the existence of chromosome-end-specific factors contributing to TL regulation.

Here, we use Nanopore sequencing to characterize the full TL distribution at each individual chromosome extremity and globally, for 100 natural yeast strains. By taking advantage of natural variations of the composition of the subtelomeric regions in these strains, we find that the presence of Y' elements in the subtelomere is associated with longer telomeres, with the notable exception of chromosome end III-L as previously reported (O'Donnell et al., 2023; Teplitz et al., 2024). We also show that inter-extremity differences contribute significantly to the spread of the global TL distribution in many strains. We thus characterize the different features underlying TL variations within and across strains.

Results

Single-molecule sequencing provides a complete view of telomere length distribution

We developed a pipeline to identify telomeric sequences in individual Nanopore reads (Fig. 1A). After base calling and adapter removal, reads are mapped onto the genome assembly of the target strain. The reads that align with chromosome extremities are scanned for telomeric sequences using Telofinder, a script designed to specifically detect the degenerated yeast telomere motif $TG_{1-3}/C_{1-3}A$ (O'Donnell et al., 2023). Finally, after filtering steps ensuring that the telomere sequences are not truncated and not internal (Supp. Fig. S1A), the length of all validated telomere sequences are aggregated into a high-resolution TL distribution (Fig. 1B-C). Importantly, because typical Nanopore reads are 30 to 200 times longer than TL, the read length did not affect the computed TL distribution (Supp. Fig. S1B), ensuring the robustness of the method with respect to variations during DNA extraction and sequencing.

First, we tested the ability of our methodology to detect biologically relevant TL differences by comparing TL distributions between the wild-type (WT) W303 laboratory strain and its *yku80Δ* derivative in which the function of the Ku complex is impaired, leading to significantly shorter telomeres (Porter et al., 1996). Our method measured a mean TL that was ~164 bp shorter in *yku80Δ* compared to WT (Fig. 1B and Supp. Fig. S1C and S1D), fully consistent with previously reported estimates (Bertuch et Lundblad 2003; Chen et al. 2018). TL distribution in *yku80Δ* mutant was also narrower (average standard deviation of 60 bp compared to 77 bp in WT), likely owing to constraints preventing telomeres from getting too short.

Second, we tested the reproducibility of our methodology by comparing TL distributions between independent sequencing experiments deriving from biological replicates. We used a natural strain isolated from pear must (AHG) for which the first set of reads was recovered from the ScRAP study (O'Donnell et al., 2023) and the second one obtained from a new Nanopore sequencing performed in the frame of this study. We found an excellent agreement between the 2 independent experiments

(Fig. 1C). We further confirmed the reproducibility of our method by sequencing biological replicates of the W303 strain and its *yku80Δ* derivative (Supp. Fig. 1C and D).

Overall, our method derives accurate, robust and biologically informative TL distributions directly from Nanopore sequencing reads. It is highly versatile, as it can be applied to any yeast sequencing dataset without particular DNA preparation, enabling the investigation of telomere length distribution diversity across numerous strains.

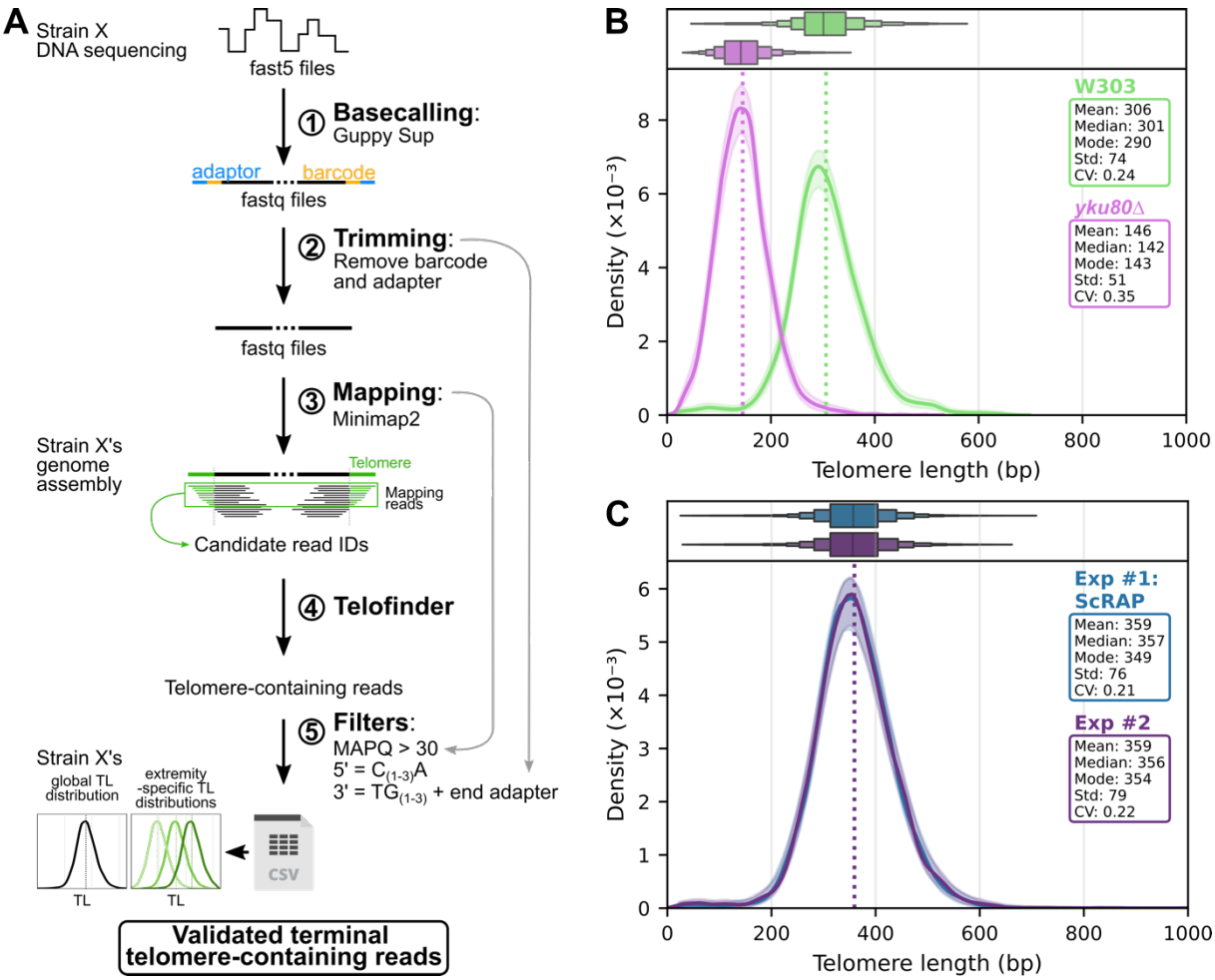

Figure 1. High-resolution TL distribution from Nanopore sequencing. (A) Bioinformatic pipeline to retrieve TL from single-molecule Nanopore sequencing. Raw electric signal from nanopores was basecalled using the super accuracy mode of Guppy (step 1). Barcodes and adapters were then removed by Porechop (step 2) and candidate sequences mapping to the telomere regions of the corresponding genome assembly were selected (step 3). The initial reads corresponding to these candidates were scanned for telomere sequences using Telofinder (step 4) and validated as telomeric reads if compliant with specific quality filters (step 5), including a mapping quality (MAPQ) threshold and test for the presence of an adapter at the 3' end of

the molecule for $TG_{(1-3)}$ sequences (grey arrows). Global and extremity-specific TL distributions are then computed. (B) TL distributions of wild-type W303 strain and its *yku80Δ* derivative, represented by a boxplot (upper graph; successive box edges indicate 50%, 75%, 87.5%, 93.75%, etc., of the data) or by its density (lower graph), obtained from Nanopore sequencing using the bioinformatic analysis pipeline described in (A). The shaded area indicates the 95% confidence interval of the sampling noise, inferred from bootstrapping. Mean value is indicated by the dotted line. Detailed metrics are shown in the insets. (C) TL distribution of strain AHG, using data from (O'Donnell et al., 2023) ("Exp #1") and from a new sequencing ("Exp #2"). Same representation as in (B).

TL distributions are highly diverse across *S. cerevisiae* strains

We measured TL distributions in a panel of 100 strains, representing the ecological and geographical diversity of *S. cerevisiae* (O'Donnell et al., 2023) (Fig. 2A and B). We obtained the species-level TL distribution by aggregating data from all telomeric reads from the 100 strains, providing the first global view of individual TL variation in *S. cerevisiae* (Supp. Fig. S2A). The mean ± standard deviation TL was measured at 359 ± 129 bp, slightly longer than commonly reported estimates in laboratory strains (Wellinger and Zakian, 2012), with the 1st and 99th TL percentiles at 117 and 810 bp, respectively.

Between the strains, the mean TL varied significantly, ranging from 176 bp in strain AFI, associated with European wine, to 682 bp in strain BGN_3a, sampled from French dairy, with ~50% of strains having TL means within the 300-400 bp range. Examples of 4 strains with diverse TL distributions in terms of means and standard deviations are shown in Fig. 2B. Within the same clade, while some strains displayed similar TL distributions (e.g., ADI and CFS, from clade "Wine European"), others exhibited vastly different ones (e.g., BAP_1a and AMM_1a, from clade "CHN-V") (Fig. 2A and Supp. Fig. S2B-D). Notably, differences between mean TLs did not correlate with phylogenetic distances, whether within or between clades (Supp. Fig. S2B). Correlation could only be detected for the most closely related strains (R^2 = 0.4 for the top 4% closest pairs of strains, Supp. Fig. S2C). Overall, this level of TL variation across strains far exceeds that observed between biological replicates (Fig. 1C, Supp. Fig. 1C and D), indicating that TL is controlled in a strain-specific manner.

Additionally, comparing TL distributions across strains revealed striking variations in others metrics, including several that were previously inaccessible. Notably, the standard deviation of the distributions varied from 44 to 228 bp, correlating with the mean (Supp. Fig. S2E). The coefficient of variation (CV, computed as the standard deviation divided by the mean) was around 0.26, but varied from 0.19 to 0.39, indicating that the standard deviation did not simply scale with mean TL (Supp. Fig. S2F). Skewness displayed values from -0.51 to 4.79, with most strains (n = 85 out of 100) exhibiting positively skewed distributions, an asymmetry consistent with the counter-selection of critically short telomeres and with short telomeres being preferentially elongated by telomerase, as supported by mathematical modelling (Xu et al., 2013). The robust kurtosis values ranged from 3.41 to 7.22, with a mean of 5.22, indicating a tendency toward leptokurtic distributions, i.e. shaped with sharper peaks and heavier tails compared to a normal distribution, suggesting a rather loose regulation of very long telomeres.

The factors that control TL distributions and their variations across strains are not known. We tested whether the ecological origins of the strains could influence TL but did not find any significant difference in mean TL, although there was a trend toward longer telomeres in domesticated compared to wild strains (Fig. 2C), consistent with recent findings (D'Angiolo et al., 2023). However, as previously suggested (O'Donnell et al., 2023), polyploids, particularly tetraploids, exhibited significantly longer mean TL than haploids and diploids, the latter two being indistinguishable (Fig. 2D).

Thus, analysis of TL distribution across a diverse panel of *S. cerevisiae* strains revealed substantial variability, not only in the mean, but also in other features of the distribution, and their genetic and/or environmental origins remain to be determined.

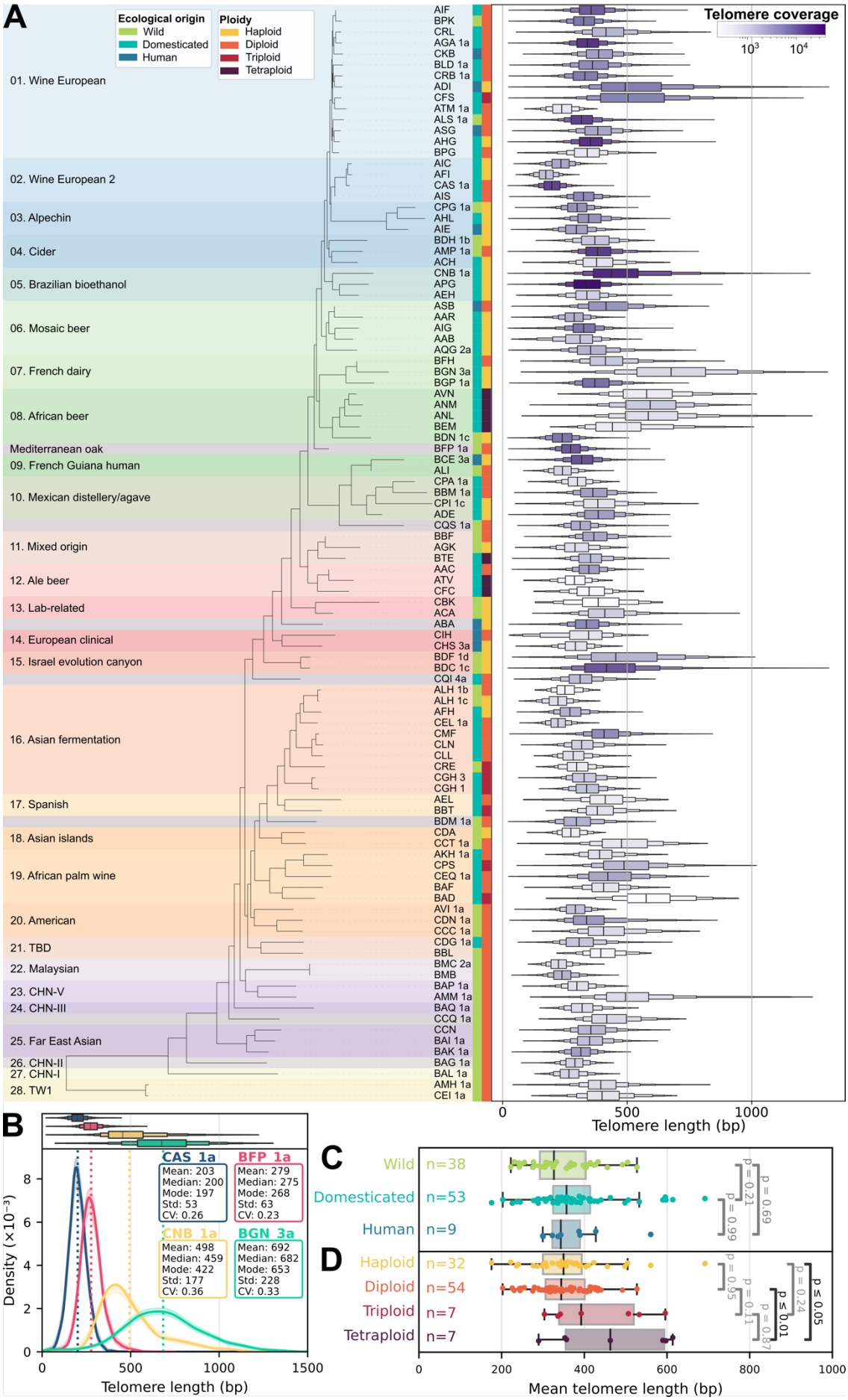
A
Ecological origin
Wild
Domesticated
Human
Ploidy
Haploid
Diploid
Triploid
Tetraploid
Telomere coverage
10^3
10^4
01. Wine European
02. Wine European 2
03. Alpechin
04. Cider
05. Brazilian bioethanol
06. Mosaic beer
07. French dairy
08. African beer
Mediterranean oak
09. French Guiana human
10. Mexican distellery/agave
11. Mixed origin
12. Ale beer
13. Lab-related
14. European clinical
15. Israel evolution canyon
16. Asian fermentation
17. Spanish
18. Asian islands
19. African palm wine
20. American
21. TBD
22. Malaysian
23. CHN-V
24. CHN-III
25. Far East Asian
26. CHN-II
27. CHN-I
28. TW1
AIF
BPK
CRL
AGA 1a
CKB
BLD 1a
CRB 1a
ADI
CFS
ATM 1a
ALS 1a
ASG
AHG
BPG
AIC
AFI
CAS 1a
AIS
CPG 1a
AHL
AIE
BDH 1b
AMP 1a
ACH
CNB 1a
APG
AEH
ASB
AAR
AIG
AAB
AQG 2a
BFH
BGN 3a
BGP 1a
AVN
ANM
ANL
BEM
BDN 1c
BFP 1a
BCE 3a
ALI
CPA 1a
BBM 1a
CPI 1c
ADE
CQS 1a
BBF
AGK
BTE
AAC
ATV
CFC
CBK
ACA
ABA
CIH
CHS 3a
BDF 1d
BDC 1c
CQI 4a
ALH 1b
ALH 1c
AFH
CEL 1a
CMF
CLN
CLL
CRE
CGH 3
CGH 1
AEL
BBT
BDM 1a
CDA
CCT 1a
AKH 1a
CPS
CEQ 1a
BAF
BAD
AVI 1a
CDN 1a
CCC 1a
CDG 1a
BBL
BMC 2a
BMB
BAP 1a
AMM 1a
BAQ 1a
CCQ 1a
CCN
BAI 1a
BAK 1a
BAG 1a
BAL 1a
AMH 1a
CEI 1a
0
500
1000
Telomere length (bp)
B
CAS 1a
Mean: 203
Median: 200
Mode: 197
Std: 53
CV: 0.26
BFP 1a
Mean: 279
Median: 275
Mode: 268
Std: 63
CV: 0.23
CNB 1a
Mean: 498
Median: 459
Mode: 422
Std: 177
CV: 0.36
BGN 3a
Mean: 692
Median: 682
Mode: 653
Std: 228
CV: 0.33
Density ($\times 10^{-3}$)
0
2
4
6
8
0
500
1000
1500
Telomere length (bp)
C
Wild
n=38
Domesticated
n=53
Human
n=9
p = 0.21
p = 0.99
p = 0.69
D
Haploid
n=32
Diploid
n=54
Triploid
n=7
Tetraploid
n=7
p = 0.65
p = 0.11
p = 0.24
p ≤ 0.01
p = 0.87
p ≤ 0.05
0
200
400
600
800
1000
Mean telomere length (bp)

Figure 2. Natural diversity of TL distribution in *S. cerevisiae*. (A) Phylogenetic tree and clade classification from (O'Donnell et al., 2023). For each strain, the TL distribution is represented as a boxplot. (B) TL distributions for 4 strains, selected to illustrate the diversity in means and standard deviations. (C) Boxplots of TL means of wild, domesticated and human-associated strains. *p*-values from Tukey's HSD tests are shown (initial ANOVA test *p*-value = 0.22). (D) Boxplots of TL means of haploid, diploid, triploid and tetraploid strains. *p*-values from Tukey's HSD tests are shown (initial ANOVA test *p*-value = 0.003). Non-significant differences are shown in grey and significant ones in black.

TL distributions are extremity-specific and inter-extremity differences vary across strains

Chromosome-end-specific differences in steady-state TL have been reported in laboratory yeast strains, using Nanopore sequencing (Sholes et al., 2022). We previously extended this observation to diverse natural yeast strains using estimates based on genomes assemblies (O'Donnell et al., 2023). These observations were enabled by the long reads allowing accurate anchoring of telomere sequences to their corresponding chromosome extremity. As a striking example, telomere III-L was found to be significantly longer than the others in many strains, except in the ones containing one or more Y' elements in this subtelomere (O'Donnell et al., 2023; Sholes et al., 2022; Teplitz et al., 2024).

Here, beyond deriving strain-specific TL distributions, our pipeline assigned each telomere-containing read to a specific chromosome end using the subtelomere sequence as an anchor, enabling us to obtain the complete distributions of extremity-specific TLs (Fig. 3A). The means of these extremity-specific TL distributions correlated well with the TL values obtained from genome assemblies ($r^2 = 0.9$) with most values differing by less than 10% (Supp. Fig. S3). However, in 12% (resp. 1%) of extremities, assembly-based TL values underestimated (resp. overestimated) the mean TL by 20% or more. In some cases, the underestimation reached up to 95% of the read-based mean TL. Importantly, the method presented in this work not only correct for the assembly-level inaccuracies at telomeres but also provides whole extremity-specific TL distributions, which were not accessible from genome assemblies and contain much richer information than simple assembly-based values.

While we confirmed that TL means were extremity-specific for most strains, we found that the extent of the inter-extremity differences within a strain varied widely (Fig. 3A). For example, strain APG associated with Brazilian bioethanol exhibited differences in mean TL between extremities while the standard deviation across extremities remained rather constant. Conversely, strain AMH_1a (clade "TW1") showed variable standard deviations but quite homogeneous TL means. In contrast, in AMM_1a (clade "CHN-V"), both TL means and standard deviations (and even CVs) were very variable from extremity to extremity. Finally, other strains displayed similar extremity-specific TL distributions both in terms of mean and standard deviation at all extremities (e.g. BAK_1a, from clade "Far East Asian"). To systematically survey the inter-extremity differences both in means and CVs (as illustrated by the 4 examples of Fig. 3A), we plotted the inter-extremity CV of the means and the inter-extremity CV of the CVs on a 2D graph for all strains (Fig. 3B). Here, the CV was chosen to measure variation preferentially over the standard deviation, because it allowed direct comparison by normalizing standard deviation by the mean. Our panel of strains was spread over the 2D space of this plot, revealing distinct TL distribution patterns across extremities. We also placed W303 and its derived *yku80Δ* mutant on this plot, showing that they displayed low inter-extremity variations. Upon closer inspection of their extremity-specific TL distributions, we found that the inter-extremity differences displayed in the wild-type W303 were suppressed in *yku80Δ*, in particular the longer TL at ChrIII_L (Fig. 3C). Overall, some strains exhibited high variability in both extremity-specific mean TL and CV, while others maintained consistent values. Some showed varying extremity-specific mean TL with a stable CV, whereas others had a stable mean but variable CVs.

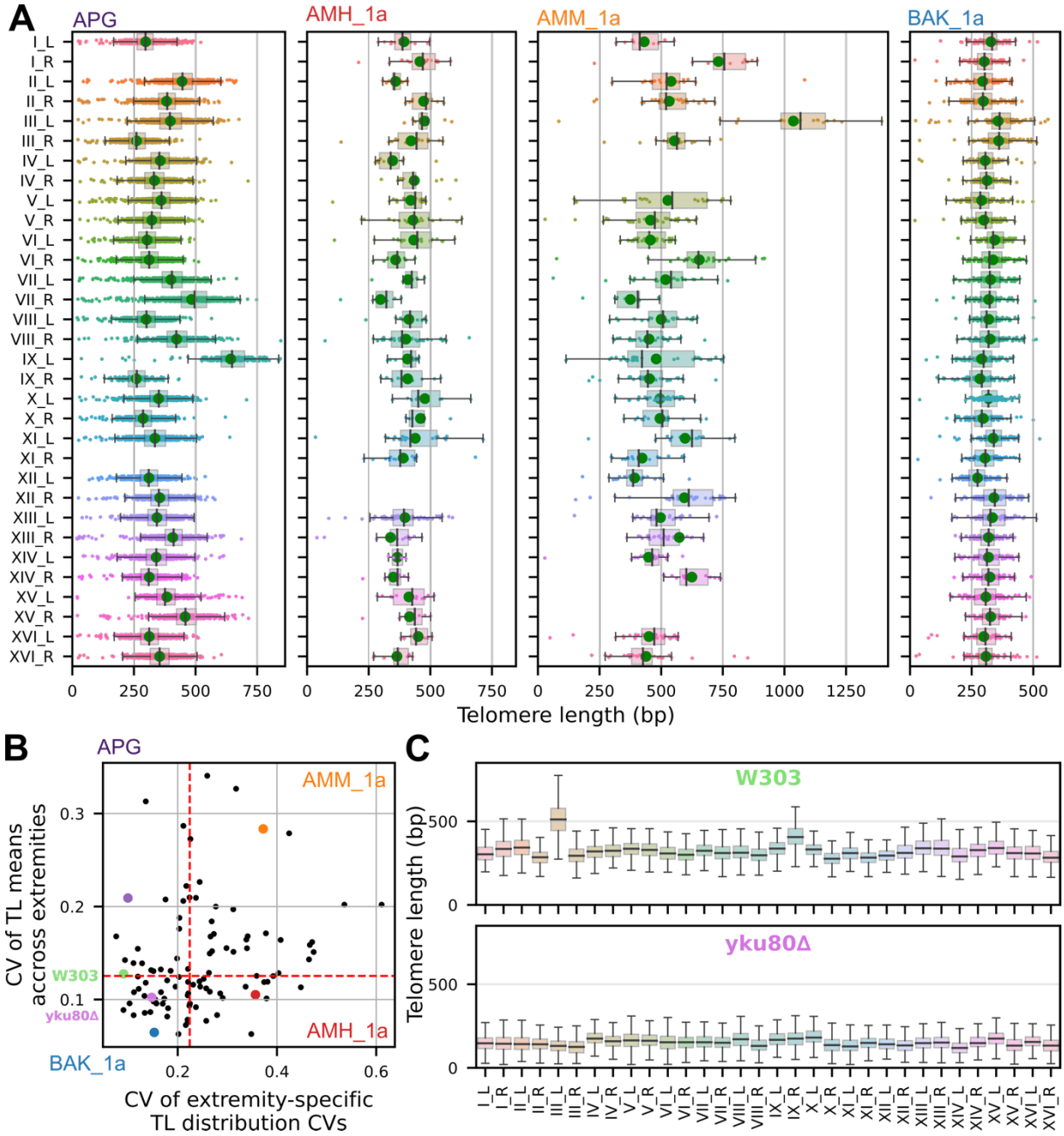

Figure 3. Diversity of extremity-specific telomere length in natural *S. cerevisiae* strains. (A) TL distributions at every chromosome extremity for 4 selected strains, represented as boxplots with individual read TL plotted as dots. Some extremities are missing in the corresponding genome assembly (O'Donnell et al., 2023), preventing TL distribution analysis. (B) 2D-plot displaying the CV of extremity-specific TL means and the CV of extremity-specific TL distributions CVs. Each strain is represented as a dot and this 2D space is divided vertically and horizontally into 2 equally sized areas. The 4 strains shown in (A) as well as W303 and its *yku80Δ* derivative are labelled in color. (C) TL distributions at every chromosome extremity for W303 and its derived *yku80Δ* mutant.

Large global TL variations primarily result from inter-extremity differences

Given the substantial inter-extremity differences in mean TL observed at the strain level, we systematically investigated their contribution to global TL variations, as compared to the contribution of within-extremity variations. Put differently, do global strain-specific TL variations primarily reflect TL variations at each extremity or are they driven by differences in mean TL between extremities?

To quantify the contribution of inter-extremity differences for each strain, we computed the CV of the extremity-specific TL means, reflecting the inter-extremity variation, and plotted it against the CV of the global TL distribution (Fig. 4A, red, and 4B). We found a strong correlation between these two values across strains (r^2 = 0.82, p-value = $4x10^{-25}$). Strikingly, the slope of the correlation was equal to 1.1, indicating that any increase in global TL variation stemmed mostly from accrued inter-extremity differences. Conversely, we plotted the mean of the CVs of extremity-specific TL distributions in a strain, reflecting the average within-extremity variation, against the CV of the global TL distribution (Fig. 4A, blue). We observed that the mean extremity-specific CV scaled with the global CV with a much lower slope of 0.3 and correlation (r^2 = 0.53, p-value = $1x10^{-8}$). Even strains with high global CVs (e.g. >0.3) still showed moderately increased extremity-specific CVs of 0.2-0.25, differing only slightly from strains with lower global CVs.

Thus, global TL distribution variation at the strain level comprises two components: the variations found at each extremity-specific TL distribution and the differences in TL between extremities. The latter accounts for most of the observed TL variations between strains.

The length of the shortest telomeres is constant at the species level

Telomerase acts preferentially on shorter telomeres and prevents them from reaching a critical length (Teixeira et al., 2004). Telomere elongation, combined with progressive telomere shortening, thus constrains the average length of the shortest telomeres in a population of cells. Given the extent of

natural TL variations that we uncovered, we wondered whether different strains exhibited different lengths for the shortest telomeres.

We thus calculated the lowest percentile of the TL distribution of each strain and compared it to the median TL (Fig. 4C). When plotted for all strains ranked according to increasing median TL, we found that the lowest percentile did not increase substantially and fluctuated around 138±45 bp, whereas the 99th percentile (longest telomeres) did. To quantitatively evaluate this observation, we calculated the Pearson's correlation value of the 1st and 99th percentiles with the median and obtained correlations of r^2 = 0.35 and 0.92, respectively (Fig. 4D). The value of the 1st percentile was thus only marginally correlated to the median TL length across strains.

To assess to which extent the length of short telomeres was poorly correlated to the median, we plotted the Pearson's correlation between the nth percentile and the median TL and found that the 2nd (resp. 3rd) percentile showed r^2 values of 0.66 (resp. 0.81), significantly lower than the plateau at 1 (Supp. Fig. S4A). Interestingly, for a haploid genome with 32 telomeres, the length of the shortest telomere should be found among the 1/32 = 3.1% lowest TL values in a distribution. The relatively weak correlation of the 1st-3rd percentiles to the median TL suggested that the length of the shortest telomere might be kept constant at the species level, regardless of the median TL.

Since extremity-specific differences in TL are observed in many strains, we wondered whether some particular chromosome extremities might bear shorter TLs more frequently than other. We thus evaluated the representation of each extremity in the individual telomeres shorter than the 1st (resp. 5th) percentile across strains and found no significant enrichment of any of them (Supp. Fig. S4B, resp. S4C), indicating than the telomere of any chromosome extremity could be among the shortest. In contrast, we found over-representation of ChrIII_L, ChrXII_R, ChrXIII_R and ChrXV_R among telomeres longer than the 99th percentile (only ChrIII_L among telomeres longer than the 95th percentile) of the distributions across strains (Supp. Fig. S4D and E). To go further into this analysis, we then asked whether the shortest extremity-specific telomere on average in a strain would be enriched among the

individual telomeres shorter than the 1st (resp. 5th) percentile (Supp. Fig. S4F, resp. G). For each strain, we ranked each chromosome extremity from shortest associated mean TL to longest associated mean TL and calculated their enrichment within the 1st (resp. 5th) percentile across strains. Even here, we barely found an enrichment for the strain-specific extremity associated with the shortest and second shortest mean TL within the 1st percentile (resp. the 5 shortest were enriched within the 5th percentile), with Log2 enrichment values below 1 (resp. 2) (Supp. Fig. S4F, resp. G). Conversely, extremities associated with greater mean TL in each strain were highly over-represented within the 99th (resp. 95th) percentile or longer with Log2 enrichment values exceeding 4 (resp. 2) (Supp. Fig. S4H, resp. I). Overall, we demonstrate that throughout the diversity of natural yeast strains, any chromosome extremity can bear the shortest telomere(s) with nearly equal frequency, an important finding to consider for the study of the functional consequences of the shortest telomere(s).

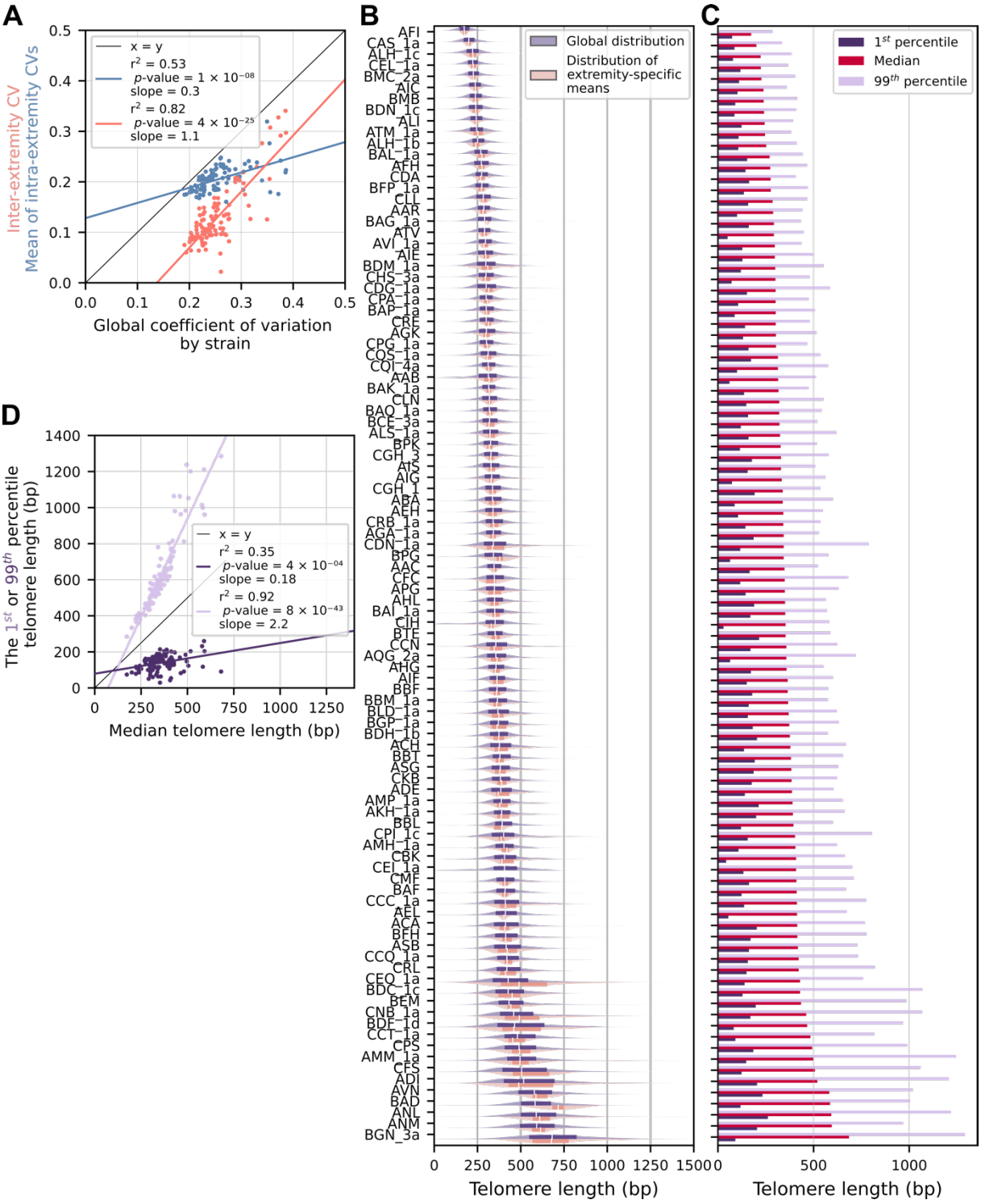

Figure 4. Contributions to global TL variations. (A) The mean of extremity-specific TL distribution CVs, reflecting the average intra-extremity variation, is plotted against global TL distribution CV as a blue dot for each strain. The CV of the distribution of extremity-specific TL means, reflecting inter-extremity variation, is plotted against global TL distribution CV as a red dot for each strain. (B) Global TL distribution (blue, upper part of each graph) and distribution of extremity-specific TL means (salmon, lower part of each graph) is displayed for all strains, ranked by increasing median TL. A boxplot representing 50% of the values is superimposed on the density plot. (C) Median, 1st and 99th percentiles of the global TL distribution are displayed

for each strain. Strains are ordered by increasing median TL. (D) The 1st (purple) and 99th (mauve) percentiles of the global TL distribution for each strain are plotted against the median TL and the linear regressions are shown.

Subtelomeric Y' elements are linked to longer telomeres in cis

The observation that TL distributions were often extremity-specific suggested the presence of cis regulatory elements. We thus compared extremity-specific TL distributions across strains, with or without elements commonly found at subtelomeres such as X and Y' elements, and Ty5 retrotransposons. The presence or absence of the X element in the adjacent subtelomere did not significantly affect the mean TL (Fig. 5A). In contrast, the presence of Ty5 in the cis subtelomere was significantly associated with longer telomeres (Fig. 5B), although the unbalance between the number of subtelomeres with and without Ty5 prevented more detailed analyses at the level of each extremity across strains.

The presence of Y' elements in the subtelomere was also associated with a significantly greater extremity-specific mean TL (383 vs 356 bp), as well as increased standard deviation and CV (Fig. 5C-E). Looking closely at each chromosome extremity across strains, the direction of this association was verified for 27 out of 32 extremities, with 3 showing statistical significance (Supp. Fig. S5A). Among the extremities with the opposite relationship between presence of Y' element and mean TL, although not statistically significant, we found ChrIII-L and ChrVII-R, for which assembly-based TL values were previously noted as longer in the absence of Y' element across strains (O'Donnell et al., 2023). Consistently, with our read-based method, ChrIII-L mean TL was significantly greater than all others when considering extremities without Y' element in the subtelomere, whereas in the presence of Y' elements, no extremity stood out (Supp. Fig. S5B-C).

Globally, the subtelomeric Y' element is associated with longer and more spread telomeres in cis, thus providing a local genomic feature as a cause for inter-extremity TL differences. Whether the Y' element

directly regulates TL through a binding factor or alters the local chromatin structure thereby affecting TL homeostasis remains to be investigated.

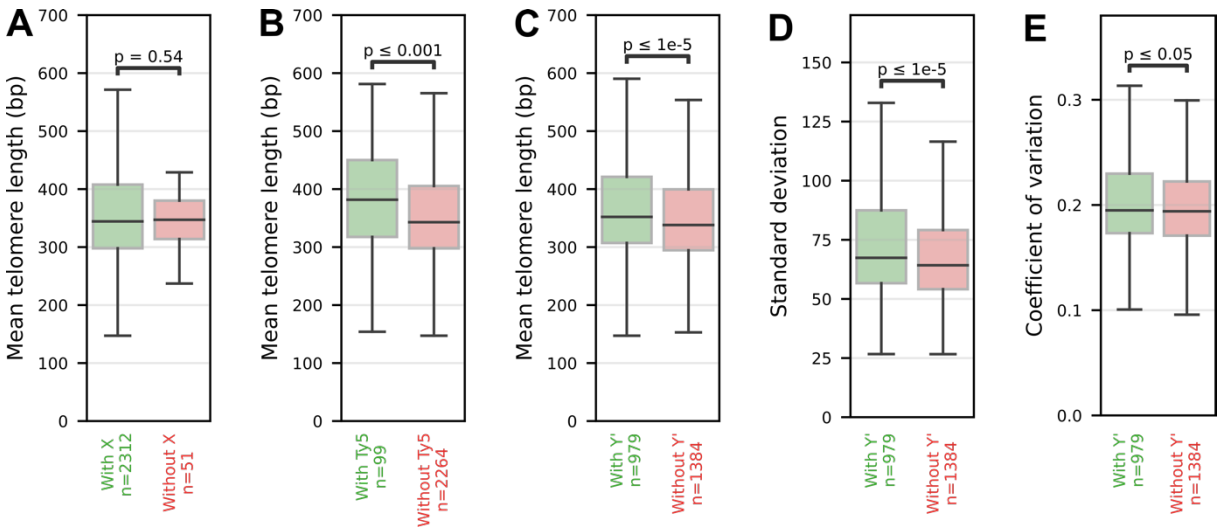

Figure 5. Association between the presence of subtelomeric elements and telomere length. (A) Distributions of mean TLs across extremities of all strains, clustered depending on the presence or absence of X element in the corresponding subtelomere. The number of chromosome extremities in each group is indicated by "n". *p*-value from a Student's t-test is indicated. (B) Same as (A) with respect to the presence or absence of Ty5 retrotransposon in the corresponding subtelomere. (C) Same as (A) with respect to the presence or absence of Y' elements in the corresponding subtelomere. (D) Same as (C) but for the standard deviation of the extremity-specific TL distribution. (E) Same as (C) but for the coefficient of variation of the extremity-specific TL distribution.

Discussion

S. cerevisiae is a major genetic model organism in the telomere field (Teixeira, 2013; Wellinger and Zakian, 2012). In contrast to the large body of works using laboratory strains, relatively little is known about TL regulation and distribution in natural yeast strains. Recent large-scale population genomics studies revealed a great genetic diversity encompassing, beyond single nucleotide polymorphisms, structural variations, aneuploidies and polyploidies (O'Donnell et al., 2023; Peter et al., 2018). We consider TL as yet another genomic variant, albeit one that is dynamic and complex, but its diversity

within the species has not yet been investigated in depth. Exploring and characterizing TL distributions in natural strains would expand our understanding of telomere regulation and help interpret telomere biology in light of the budding yeast natural history, ecology and evolution. As a first step towards this goal, in this work, we provide a detailed characterization of TL distributions in a panel of natural isolates recently sequenced by long-read Nanopore technology (O'Donnell et al., 2023) and derive biologically relevant insights.

We evidence a large diversity of TL distributions across a representative panel of strains. Compared to the commonly studied laboratory strains, some strains show much shorter or longer mean TL, which might be explained by genetic variants and differences in the molecular mechanisms governing TL regulation. Genome-wide association studies (GWAS) or linkage analyses could leverage the genetic diversity to identify new genes or variants contributing to TL regulation. Furthermore, TL diversity goes beyond variations of the mean. The standard deviation, CV, skewness and kurtosis all show substantial differences, even within the same clade. Because TL distribution is fundamentally shaped by the dynamics of telomere shortening and elongation by telomerase, the diversity of TL distributions in natural strains suggests that they differ in some or all aspects of these mechanisms. Our work therefore represents a useful resource for future investigations into TL regulation. From an evolutionary perspective, our observations suggest that TL distribution might not be the object of strong selective pressure, except perhaps on short telomeres. Alternatively, TL distribution could be a fast-evolving trait associated with some fitness phenotypes dependent on the ecological environment that remain to be discovered.

Since we observed extremity-specific TL means, we leveraged the structural diversity found at the subtelomeres of the ScRAP panel of strains to interrogate the contribution of subtelomeric elements in cis to TL regulation. We detected a small but significant difference in mean TL between extremities with or without Y' elements across all analyzed strains (difference of 27 bp, 95% confidence interval of ±10 bp). Remarkably, the direction of this difference, i.e. longer telomeres at Y'-containing extremities,

was verified in 27 out of 32 extremities, arguing for a general effect of Y' elements on cis TL. That being said, the extent of the difference varied from one extremity to another and at some of them, the relationship was reversed, suggesting that the local context can modulate the role of the Y' element in TL regulation. For ChrIII-L specifically, Teplitz and colleagues proposed a molecular mechanism, involving heterochromatin on the HML locus promoting telomerase recruitment in the absence of Y' element at the ChrIII-L subtelomere, to explain the longer telomeres observed without Y' element at this extremity (Teplitz et al., 2024). To our knowledge, except for ChrIII-L, there was no previous evidence of subtelomeric elements directly affecting TL in cis. Deletion of all X elements, Y' elements or both in a strain engineered to contain only 3 chromosomes showed no significant effect on TL as measured by Southern blot (Hu et al., 2024), although this technique might not have the resolution to detect subtle differences. Interestingly, differences in TL regulation between X-only and Y'-containing subtelomeres could be revealed in the *tel1Δ* mutant, in which the preferential elongation of short telomeres is lost (Arneric and Lingner, 2007; Craven and Petes, 1999), suggesting an indirect or buffered effect of subtelomere structure on TL. We find that inter-extremity TL differences are eliminated in the *yku80Δ* mutant compared to wild-type W303, implicating the Ku complex in mediating TL regulation in cis. We thus propose that, as a general rule, Y' elements and possibly Ty5 retrotransposons, but not X elements, positively influence TL in cis, thus contributing to TL variations between extremities. Further investigation is needed to uncover the associated molecular mechanisms.

While yeast cells constitutively express telomerase, the signaling pathway triggering replicative senescence if a telomere becomes critically short is functional, as revealed in telomerase mutants (Lundblad and Szostak, 1989). Selection against critically short telomeres could therefore shape the regulation of TL so that, with active telomerase, the steady-state TL distribution would not allow senescence to occur at a significant frequency. Conversely, cells might incur a fitness cost for maintaining long telomeres, because replication forks slow down at telomeres and can be stalled by Rap1 and secondary structures such as G-quadruplexes (Douglas and Diffley, 2021; Ivessa et al., 2002;

Lopes et al., 2011; Paeschke et al., 2011), although no fitness defect was detected in a lab strain carrying longer telomeres (Harari et al., 2017). Keeping these two plausible selection forces in mind, our finding that the shortest 1-3% telomeres in all strains are maintained at a roughly similar length, with only a relatively low correlation with the median length, would be consistent with selection acting only on the shortest telomeres so as to keep them just above a detrimental threshold. The 1st percentile measured at ~138 bp would provide some margin for the shortest telomeres to shorten and be elongated without crossing proposed critical threshold lengths between 20 and 80 bp (Abdallah et al., 2009; Kockler et al., 2021; Rat et al., 2025; Sholes et al., 2022). Accordingly, we show that natural strains with much longer telomeres generally display steady-state distributions that are more spread, hence the observed large CVs, rather than tight distributions with increased means. Remarkably, the shortest telomeres in a strain do not belong preferentially to one or a few defined extremities, but rather, telomeres from any extremity have a similar probability to be among the shortest ones, thus spreading the potential risk associated with being short.

To conclude, we combine the most resolutive TL measurement method to date with a large sample set representing the phylogenetic and ecological diversity of *S. cerevisiae* to characterize the variability of TL distributions, its sources, and infer biological insights into TL regulation.

Material and Methods

Strains and sequencing data

We used data from 100 sequenced strains presented in (O'Donnell et al., 2023) (Supp. Table S1). In addition, strain AHG was sequenced again to test the reproducibility of our method, as well as WT W303 (*Mat***a** *ura3-1 trp1-1 leu2-3,112 his3-11,15 can1-100 RAD5 ADE2*) and *yku80Δ* (strain name: yZX440) strains. *YKU80* was deleted by replacement with *LEU2* using plasmid pFA6a-*LEU2* (Houseley and Tollervey, 2011) and standard PCR-based replacement methods (Longtine et al., 1998). Strains were grown in yeast extract, peptone, dextrose media (YPD) at 30°C.

Genomic DNA extraction

We grew yeast cells picked from a single colony in 10 mL YPD media at 30 °C overnight. A total number of cells of 7×10^9 or less was used for DNA extraction. High molecular weight (HMW) DNA was extracted using QIAGEN Genomic-tip 100/G according to the manufacturer's instructions. Alternatively, genomic DNA was extracted using the CTAB method, as described in (Kaur et al., 2018). Briefly, 2×10^9 cells were incubated for 20 min at 37°C in spheroblasting solution (1 M sorbitol, 50 mM KPO_4, 10 mM EDTA) with 1.25 mg of zymolyase 20T. Spheroblast were pelleted and resuspended in CTAB extraction solution (3% CTAB, 100 mM Tris-HCl, 25 mM EDTA, 2 M NaCl and 2% PVP40) with 12.5 µL of 20 mg/mL proteinase K and 2 µL of 100 mg/mL RNAse A. After 40 min of incubation at 37°C, cellular debris were separated from DNA using chloroform:isoamyl alcohol solution. Then the CTAB-DNA mixture was precipitated using CTAB dilution solution (1% CTAB, 50 mM Tris-HCl, 10 mM EDTA). The precipitate was washed with NaCl solution 1 (0.4 M NaCl, 10 mM Tris-HCl, 1 mM EDTA) to remove the excess of CTAB and then resuspended in NaCl solution 2 (1.42 M NaCl, 10 mM Tis-HCl, 1 mM EDTA). DNA was precipitated in adding 2 volumes of absolute ethanol and centrifugation for 2 min at 12,000

rpm. After a wash with 70% ethanol, DNA were resuspended in TE buffer. DNA quantity was controlled by Qubit dsDNA HS Assay.

Nanopore sequencing

Library preparation and ONT sequencing were performed based on the protocol of '1D Native barcoding genomic DNA with EXP-NBD104 and SQK-LSK108' when using FLO-MIN106 MinION flow cells and protocol of '1D Genomic DNA by Ligation with EXP-NBD104 and SQK-LSK109—PromethION' when using the V2 FLO-PRO002 flow cell. DNA repair and end preparation were performed on up to 2 µg of HMW DNA per sample using the NEBNext FFPE DNA Repair Mix with the following reaction setup: 48 µL DNA, 3.5 µL NEBNext FFPE DNA Repair Buffer, 2 µL NEBNext FFPE DNA Repair Mix, 3.5 µL Ultra II End Prep Reaction Buffer and 3 µL Ultra II End Prep Enzyme Mix, incubated at 20 °C for 15 min followed by 65 °C for 15 min. DNA size selection was then carried out using AMPure XP Beads (1:1 ratio) followed by native barcode ligation (22.5 µL DNA, 2.5 µL native barcode provided by EXP-NBD104 kit and 25 µL Blunt/TA Ligase Master Mix; at 25 °C for 20 min). After another round of AMPure XP bead clean-up (1:1 ratio), the samples were pooled together and adaptors were ligated at 25 °C for 15 min (65 µL DNA, 5 µL adapter mix II (AMII) provided by EXP-NBD 104 kit, 20 µL NEBNext Quick Ligation Reaction Buffer and 10 µL Quick T4 DNA Ligase; at 25 °C for 15 min). The adaptor-ligated DNA sample was cleaned up by adding a 0.4× volume of AMPure XP beads followed by incubation for 5 min at room temperature. When using the SQK-LSK108 kit for FLO-MIN106 MinION flow cells, two successive 140-µL adapter bead binding buffer (ABB) washes were performed. When using the SQK-LSK109 kit for FLO-PRO002 flow cells, two successive 250-µL L fragment buffer washes were performed. The final library was eluted in 15 µL Elution Buffer and loaded into the MinION or PromethION flow cells according to the ONT manuals.

Telomere length measurement pipeline

All bioinformatic analyses and pipelines were coded in Python (version 3.9.12) with Numpy (version 1.21.6) and Pandas (version 1.4.2) libraries installed. Raw fast5 files generated from sequencing were basecalled using the Super Accurate model of Guppy (version 6.4.2) followed by Porechop (version 0.2.4; https://github.com/rrwick/Porechop) removal of adapters and barcodes (Fig. 1A). Reads were then mapped to the genome assembly of the same strain using minimap2 (https://github.com/lh3/minimap2) (Li, 2018) and the ones mapping to telomeric regions with MAPQ ≥ 30 were retrieved. We then ran Telofinder (https://github.com/GillesFischerSorbonne/telofinder) (O'Donnell et al., 2023) on these selected reads to detect the terminal telomeric sequences. Telofinder calculates sequence entropy and proportions of dinucleotides (TG/CA, GG/CC and GT/AC) within a 20-bp sliding window, enabling the detection of the degenerated yeast telomere motifs $(TG)_{1\text{-}6}TG_{2\text{-}3}/C_{2\text{-}3}A(CA)_{1\text{-}6}$ on Nanopore reads when specific thresholds are met. To ensure that the identified sequences corresponded to true telomeres, we verified that a telomeric sequence found at the beginning (5'-end) of a read was C-rich, and one found at the end (3'-end) was G-rich, as expected based on the 5'-3' orientation of sequencing through the pores. Additionally, for G-rich telomere sequences found at 3' ends, we only selected reads that contained a terminal adapter to ensure that they were not truncated before sequencing reached the end of the DNA molecule.

Phylogenetic tree processing and visualization

The phylogenetic tree (Fig. 2A) was built based on data from (O'Donnell et al., 2023), with adjustments to retain only the 100 leaves representing the 100 strains analyzed here. We used the Bio.Phylo package from Biopython (v1.79) for phylogenetic analysis and iTOL v6 (Letunic and Bork, 2024) for visualization.

Plots and statistical analyses

Python with Seaborn (version 0.11.2) and SciPy (version 1.7.3) libraries was used to draw plots and perform statistical analyses.

Data availability

All new sequencing data (fast5, fastq and fasta files) were deposited in the European Nucleotide Archive (https://www.ebi.ac.uk/ena/ browser/home) under the project PRJEB89249.

Code availability

All scripts used in this publication will be available at https://github.com/Telomere-Genome-Stability/Yeast2025.

List of Supplementary Figures and Table

Supplementary Figure 1. Robust TL distribution measurement from Nanopore reads.

Supplementary Figure 2. Characterization of TL distribution metrics in natural strains.

Supplementary Figure 3. Comparison between assembly-derived and read-derived extremity-specific TL values.

Supplementary Figure 4. Contributions of specific extremities to the shortest or longest telomeres.

Supplementary Figure 5. The presence of a Y' element is correlated with longer TL.

Supplementary Table S1. List of strains and accessions.

Acknowledgement

We thank members of GF's and ZX's laboratories for discussions. Research in ZX's laboratory was supported by Agence Nationale de la Recherche (ANR-24-CE12-7740-01 and ANR-24-CE12-0083-03) and Institut Nationale du Cancer (PLBIO20-312). Research in GF's laboratory was supported by Agence Nationale de la Recherche (ANR-24-CE12-7740-02). EK's work at the Biomics Platform, C2RT, Institut Pasteur, Paris, France, was supported by France Génomique (ANR-10-INBS-09).

Author contributions

CG: Methodology, Investigation, Software, Formal Analysis, Writing – Original Draft. EK: Methodology, Investigation, Software. NA: Investigation. OI: Investigation. GF: Conceptualization, Methodology, Supervision, Writing – Original Draft. ZX: Conceptualization, Supervision, Writing – Original Draft. All authors: Writing – Review & Editing.

Conflicts of interest

The authors declare no conflict of interest.

References

Abdallah, P., Luciano, P., Runge, K.W., Lisby, M., Geli, V., Gilson, E., and Teixeira, M.T. (2009). A two-step model for senescence triggered by a single critically short telomere. Nat Cell Biol *11*, 988-993.

Allsopp, R.C., Vaziri, H., Patterson, C., Goldstein, S., Younglai, E.V., Futcher, A.B., Greider, C.W., and Harley, C.B. (1992). Telomere length predicts replicative capacity of human fibroblasts. Proc. Natl. Acad. Sci. U. S. A. *89*, 10114-10118.

Arneric, M., and Lingner, J. (2007). Tel1 kinase and subtelomere-bound Tbf1 mediate preferential elongation of short telomeres by telomerase in yeast. EMBO Rep *8*, 1080-1085.

Blackburn, E.H., and Chiou, S.S. (1981). Non-nucleosomal packaging of a tandemly repeated DNA sequence at termini of extrachromosomal DNA coding for rRNA in Tetrahymena. Proc. Natl. Acad. Sci. U. S. A. *78*, 2263-2267.

Craven, R.J., and Petes, T.D. (1999). Dependence of the regulation of telomere length on the type of subtelomeric repeat in the yeast Saccharomyces cerevisiae. Genetics *152*, 1531-1541.

Cristofari, G., and Lingner, J. (2006). Telomere length homeostasis requires that telomerase levels are limiting. EMBO J.

D'Angiolo, M., Yue, J.X., De Chiara, M., Barre, B.P., Giraud Panis, M.J., Gilson, E., and Liti, G. (2023). Telomeres are shorter in wild Saccharomyces cerevisiae isolates than in domesticated ones. Genetics *223*.

de Lange, T. (2018). Shelterin-Mediated Telomere Protection. Annu. Rev. Genet. *52*, 223-247.

Douglas, M.E., and Diffley, J.F.X. (2021). Budding yeast Rap1, but not telomeric DNA, is inhibitory for multiple stages of DNA replication in vitro. Nucleic Acids Res *49*, 5671-5683.

Eberhard, S., Valuchova, S., Ravat, J., Fulnecek, J., Jolivet, P., Bujaldon, S., Lemaire, S.D., Wollman, F.A., Teixeira, M.T., Riha, K.*, et al.* (2019). Molecular characterization of Chlamydomonas reinhardtii telomeres and telomerase mutants. Life science alliance *2*.

Enomoto, S., Glowczewski, L., and Berman, J. (2002). MEC3, MEC1, and DDC2 Are Essential Components of a Telomere Checkpoint Pathway Required for Cell Cycle Arrest during Senescence in Saccharomyces cerevisiae. Mol. Biol. Cell. *13*, 2626-2638.

Fajkus, J., Kovarik, A., Kralovics, R., and Bezdek, M. (1995). Organization of telomeric and subtelomeric chromatin in the higher plant Nicotiana tabacum. Mol. Gen. Genet. *247*, 633-638.

Fulcher, N., Teubenbacher, A., Kerdaffrec, E., Farlow, A., Nordborg, M., and Riha, K. (2015). Genetic architecture of natural variation of telomere length in Arabidopsis thaliana. Genetics *199*, 625-635.

Gallardo, F., Laterreur, N., Cusanelli, E., Ouenzar, F., Querido, E., Wellinger, R.J., and Chartrand, P. (2011). Live cell imaging of telomerase RNA dynamics reveals cell cycle-dependent clustering of telomerase at elongating telomeres. Mol. Cell *44*, 819-827.

Gatbonton, T., Imbesi, M., Nelson, M., Akey, J.M., Ruderfer, D.M., Kruglyak, L., Simon, J.A., and Bedalov, A. (2006). Telomere Length as a Quantitative Trait: Genome-Wide Survey and Genetic Mapping of Telomere Length-Control Genes in Yeast. PLoS Genet *2*, e35.

Gomes, N.M., Ryder, O.A., Houck, M.L., Charter, S.J., Walker, W., Forsyth, N.R., Austad, S.N., Venditti, C., Pagel, M., Shay, J.W.*, et al.* (2011). Comparative Biology of Mammalian Telomeres: Hypotheses on Ancestral States and the Roles of Telomeres in Longevity Determination. Aging Cell.

Greider, C.W., and Blackburn, E.H. (1985). Identification of a specific telomere terminal transferase activity in Tetrahymena extracts. Cell *43*, 405-413.

Harari, Y., Zadok-Laviel, S., and Kupiec, M. (2017). Long Telomeres Do Not Affect Cellular Fitness in Yeast. mBio *8*.

Houseley, J., and Tollervey, D. (2011). Repeat expansion in the budding yeast ribosomal DNA can occur independently of the canonical homologous recombination machinery. Nucleic Acids Res *39*, 8778-8791.

Hu, C., Zhu, X.T., He, M.H., Shao, Y., Qin, Z., Wu, Z.J., and Zhou, J.Q. (2024). Elimination of subtelomeric repeat sequences exerts little effect on telomere essential functions in Saccharomyces cerevisiae. eLife *12*.

Ijpma, A.S., and Greider, C.W. (2003). Short Telomeres Induce a DNA Damage Response in Saccharomyces cerevisiae. Mol. Biol. Cell *14*, 987-1001.

Ivessa, A.S., Zhou, J.Q., Schulz, V.P., Monson, E.K., and Zakian, V.A. (2002). Saccharomyces Rrm3p, a 5' to 3' DNA helicase that promotes replication fork progression through telomeric and subtelomeric DNA. Genes Dev. *16*, 1383-1396.

Jain, D., and Cooper, J.P. (2010). Telomeric strategies: means to an end. Annual Review of Genetics *44*, 243-269.

Karimian, K., Groot, A., Huso, V., Kahidi, R., Tan, K.T., Sholes, S., Keener, R., McDyer, J.F., Alder, J.K., Li, H.*, et al.* (2024). Human telomere length is chromosome end-specific and conserved across individuals. Science *384*, 533-539.

Kaur, H., Ahuja, J.S., and Lichten, M. (2018). Methods for Controlled Protein Depletion to Study Protein Function during Meiosis. Methods Enzymol. *601*, 331-357.

Kockler, Z.W., Comeron, J.M., and Malkova, A. (2021). A unified alternative telomere-lengthening pathway in yeast survivor cells. Mol. Cell.

Letunic, I., and Bork, P. (2024). Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. Nucleic Acids Res *52*, W78-W82.

Levy, D.L., and Blackburn, E.H. (2004). Counting of Rif1p and Rif2p on Saccharomyces cerevisiae telomeres regulates telomere length. Mol. Cell. Biol. *24*, 10857-10867.

Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics *34*, 3094-3100.

Liti, G., Haricharan, S., Cubillos, F.A., Tierney, A.L., Sharp, S., Bertuch, A.A., Parts, L., Bailes, E., and Louis, E.J. (2009). Segregating YKU80 and TLC1 alleles underlying natural variation in telomere properties in wild yeast. PLoS Genet *5*, e1000659.

Longtine, M.S., McKenzie, A.I., Demarini, D.J., Shah, N.G., Wach, A., Brachat, A., Philippsen, P., and Pringle, J.R. (1998). Additional modules for versatile and economical PCR-based gene deletion and modification in Saccharomyces cerevisiae. Yeast *14*, 953-961.

Lopes, J., Piazza, A., Bermejo, R., Kriegsman, B., Colosio, A., Teulade-Fichou, M.P., Foiani, M., and Nicolas, A. (2011). G-quadruplex-induced instability during leading-strand replication. EMBO J. *30*, 4033-4046.

Lundblad, V., and Szostak, J.W. (1989). A mutant with a defect in telomere elongation leads to senescence in yeast. Cell *57*, 633-643.

Marcand, S., Gilson, E., and Shore, D. (1997). A protein-counting mechanism for telomere length regulation in yeast. Science *275*, 986-990.

Monaghan, P., Eisenberg, D.T.A., Harrington, L., and Nussey, D. (2018). Understanding diversity in telomere dynamics. Philos. Trans. R. Soc. Lond. B. Biol. Sci. *373*.

Moyzis, R.K., Buckingham, J.M., Cram, L.S., Dani, M., Deaven, L.L., Jones, M.D., Meyne, J., Ratliff, R.L., and Wu, J.R. (1988). A highly conserved repetitive DNA sequence, (TTAGGG)n, present at the telomeres of human chromosomes. Proc. Natl. Acad. Sci. U. S. A. *85*, 6622-6626.

Mozdy, A.D., and Cech, T.R. (2006). Low abundance of telomerase in yeast: implications for telomerase haploinsufficiency. RNA *12*, 1721-1737.

O'Donnell, S., Yue, J.X., Saada, O.A., Agier, N., Caradec, C., Cokelaer, T., De Chiara, M., Delmas, S., Dutreux, F., Fournier, T.*, et al.* (2023). Telomere-to-telomere assemblies of 142 strains characterize the genome structural landscape in Saccharomyces cerevisiae. Nat. Genet. *55*, 1390-1399.

Paeschke, K., Capra, J.A., and Zakian, V.A. (2011). DNA Replication through G-Quadruplex Motifs Is Promoted by the Saccharomyces cerevisiae Pif1 DNA Helicase. Cell *145*, 678-691.

Peter, J., De Chiara, M., Friedrich, A., Yue, J.X., Pflieger, D., Bergstrom, A., Sigwalt, A., Barre, B., Freel, K., Llored, A.*, et al.* (2018). Genome evolution across 1,011 Saccharomyces cerevisiae isolates. Nature *556*, 339-344.

Porter, S.E., Greenwell, P.W., Ritchie, K.B., and Petes, T.D. (1996). The DNA-binding protein Hdf1p (a putative Ku homologue) is required for maintaining normal telomere length in Saccharomyces cerevisiae. Nucleic Acids Res *24*, 582-585.

Rat, A., Martinez Fernandez, V., Doumic, M., Teixeira, M.T., and Xu, Z. (2025). Mathematical model linking telomeres to senescence in Saccharomyces cerevisiae reveals cell lineage versus population dynamics. Nat Commun *16*, 1024.

Sanchez, S.E., Gu, Y., Wang, Y., Golla, A., Martin, A., Shomali, W., Hockemeyer, D., Savage, S.A., and Artandi, S.E. (2024). Digital telomere measurement by long-read sequencing distinguishes healthy aging from disease. Nat Commun *15*, 5148.

Schmidt, T.T., Tyer, C., Rughani, P., Haggblom, C., Jones, J.R., Dai, X., Frazer, K.A., Gage, F.H., Juul, S., Hickey, S.*, et al.* (2024). High resolution long-read telomere sequencing reveals dynamic mechanisms in aging and cancer. Nat Commun *15*, 5149.

Shakirov, E.V., and Shippen, D.E. (2004). Length regulation and dynamics of individual telomere tracts in wild-type Arabidopsis. Plant Cell *16*, 1959-1967.

Shampay, J., and Blackburn, E.H. (1988). Generation of telomere-length heterogeneity in Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. U. S. A. *85*, 534-538.

Sholes, S.L., Karimian, K., Gershman, A., Kelly, T.J., Timp, W., and Greider, C.W. (2022). Chromosome-specific telomere lengths and the minimal functional telomere revealed by nanopore sequencing. Genome Res. *32*, 616-628.

Teixeira, M.T. (2013). Saccharomyces cerevisiae as a Model to Study Replicative Senescence Triggered by Telomere Shortening. Front Oncol *3*, 101.

Teixeira, M.T., Arneric, M., Sperisen, P., and Lingner, J. (2004). Telomere Length Homeostasis Is Achieved via a Switch between Telomerase- Extendible and -Nonextendible States. Cell *117*, 323-335.

Teplitz, G.M., Pasquier, E., Bonnell, E., De Laurentiis, E., Bartle, L., Lucier, J.F., Sholes, S., Greider, C.W., and Wellinger, R.J. (2024). A mechanism for telomere-specific telomere length regulation. bioRxiv.

Walmsley, R.M., and Petes, T.D. (1985). Genetic control of chromosome length in yeast. Proc. Natl. Acad. Sci. U. S. A. *82*, 506-510.

Wellinger, R.J., and Zakian, V.A. (2012). Everything you ever wanted to know about Saccharomyces cerevisiae telomeres: beginning to end. Genetics *191*, 1073-1105.

Wicky, C., Villeneuve, A.M., Lauper, N., Codourey, L., Tobler, H., and Muller, F. (1996). Telomeric repeats (TTAGGC)n are sufficient for chromosome capping function in Caenorhabditis elegans. Proc. Natl. Acad. Sci. U. S. A. *93*, 8983-8988.

Xu, Z., Dao Duc, K., Holcman, D., and Teixeira, M.T. (2013). The length of the shortest telomere as the major determinant of the onset of replicative senescence. Genetics *194*, 847-857.